DOCUMENT RESUME

ED 403 314 TM 026 049

AUTHOR Gould, R. Bruce; And Others
TITLE The Development, Validation, and Applicability of "The Program Evaluation Standards: How To Assess Evaluations of Educational Programs."
PUB DATE Aug 95
NOTE 26p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Evaluation Methods; Formative Evaluation; Program Development; *Program Evaluation; *Standards; Summative Evaluation; *Test Construction
IDENTIFIERS *Joint Committee on Standards for Educ Evaluation

ABSTRACT

The work done by the Validation Panel that was commissioned by the Joint Committee on Standards for Educational Evaluation (Joint Committee) to monitor the development of "The Program Evaluation Standards: How To Assess Evaluations of Educational Programs" is described, and its conclusions summarized. This report focuses on the development process, the assumptions underlying the effort, and the applicability of the "Standards" in different contexts. Revision of the "Standards" had begun at the Joint Committee's 1990 meeting. An early decision was made to include a standard for meta-evaluation that required that the evaluation itself be formatively and summatively evaluated. True to this new standard, the Joint Committee commissioned a Validation Panel to perform that meta-evaluation function during the development of the revised "Standards." The developed "Standards" consist of 30 specific standards grouped into categories of utility, feasibility, propriety, and accuracy. Although no explicit statements of guiding assumptions are included in the "Standards," a number of implicit assumptions center on the need for educational program evaluation standards and the possibility of agreement about such standards. Representatives of the 15 organizations that comprise the Joint Committee considered the results of expert commentary, testimony at public hearings, and field tests in approving the development process. The position is taken that the development of the "Standards" was very systematic and open, and likely resulted in a set of standards that represent the state of the art in educational program evaluation. (Contains nine references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
R. BRUCE GOULD
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

ED 403 314

The Development, Validation, and Applicability of

The Program Evaluation Standards:

How to Assess Evaluations of Educational Programs

Report of the Validation Panel Commissioned by the

Joint Committee on Standards for Educational Evaluation

R. Bruce Gould, Air Force Armstrong Laboratory, Chair

David J. Basarab, Motorola University

Christine McGuire, University of Illinois

Phillip C. Robinson, PR-2 Associates

F. LeRoy Walser, University of Oklahoma

Alexandra K. Wigdor, National Academy of Sciences

TM026049

BEST COPY AVAILABLE

The Development, Validation, and Applicability of

The Program Evaluation Standards:

How to Assess Evaluations of Educational Programs

INTRODUCTION

The purpose of this report is to summarize the work and conclusions of the Validation Panel which was commissioned to monitor development of *The Program Evaluation Standards: How to Assess Evaluations of Educational Programs*, 2nd Edition, by the Joint Committee on Standards for Educational Evaluation (subsequently referred to as the Joint Committee or Committee). In keeping with the charge to the Validation Panel, this report focuses on the developmental process, the assumptions that underlie the effort, and the applicability of the *Standards* in different contexts.

BACKGROUND

Standards Projects

The Joint Committee was founded in 1975 on the recommendation of a committee composed of members of the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME) who completed the 1975 version of the *Standards for Educational and Psychological Tests and Manuals*. The AERA/APA/NCME Committee on Test Standards considered including a section on program evaluation but decided that the inclusion would not be within the scope of their task and recommended a new committee

be appointed to address the issue. A planning committee of the three organizations recommended that a separate project be established to develop standards for educational evaluation and invited other major organizations concerned with education to join the effort.

The new Joint Committee on Standards for Educational Evaluation began its work in the Fall of 1975. They initially found that there were no adequate standards for educational evaluation and believed that a set of professional standards could play a vital role in upgrading the practice of educational evaluation. Further,

> the Joint Committee foresaw several benefits from the development of sound standards: a common language to facilitate communication and collaboration in evaluation; a set of general rules for dealing with a variety of specific evaluation problems; a conceptual framework by which to study the often-confusing world of evaluation; a set of working definitions to guide research and development on the evaluation process; a public statement of the state of the art in educational evaluation; a basis for self regulation and accountability by professional evaluators; and an aid to developing public credibility for the educational evaluation field (The Joint Committee; 1981).

The Joint Committee's efforts led to the 1981 publication of the *Standards for Evaluations of Educational Programs, Projects, and Materials*. As the project was nearing completion, members of the Joint Committee concluded that there was a need to monitor and foster the use of the standards that had been developed, revise them as appropriate, and develop standards dealing with other types of evaluation in education. Therefore, the Joint Committee became a separate incorporated organization consisting of a maximum of eighteen members who are appointed either as representatives of

sponsoring organizations or members-at-large by the Joint Committee, plus a Chairperson. "The Joint Committee is incorporated exclusively for setting evaluation standards and promoting their use" (Joint Committee By-Laws, Article II, Section 1, *MISSION STATEMENT*).

The Joint Committee's second major project was begun in early 1985 to improve the quality of personnel evaluation systems in education. This resulted in the 1988 publication of *The Personnel Evaluation Standards: How to Assess Systems for Evaluating Educators*. The developmental process for both the above standards was similar and will be covered in some detail later. What was unique in the *Personnel Evaluation Standards* developmental process was that an independent Validation Panel was commissioned to examine the development and final product and provide both formative and summative evaluation information focused on the assumptions, validation procedures, and applicability of those *Standards* (Linn, Buchmann, Gould, Kellaghan, Lawrence, Robinson, & Zirkel, 1988).

Revision of the Program Evaluation Standards

At its 1989 annual meeting, the Joint Committee decided to revise the 1981 *Standards for Evaluations of Educational Programs, Projects, and Materials* because they did not include new developments in evaluation (e.g. qualitative and meta evaluation and views of validity of information emphasis shift from content, criterion, and construct validity to focus more on the gathering, interpretation, and use of the information) that had been introduced since the *Standards* were first published. Also, over 20,000 copies

of the *Standards* had been sold and many users provided oral or written comments on their use and desired changes. There was also a growing interest in using the *Standards* in areas beyond the target K-12 school settings of the 1981 *Standards*.

Revision of the *Standards* began at the Joint Committee's 1990 meeting. Focus at the meeting was to make decisions about the organization and substance of the standards and draft initial revisions of the standards themselves so that a subsequent volunteer panel of writers could critique the drafts and begin writing alternative illustrative cases. An early decision was made to include a new metaevaluation standard which said that "the evaluation itself should be formatively and summatively evaluated against these and other pertinent standards, so that its conduct is appropriately guided and, on completion, stakeholders can closely examine its strengths and weaknesses."

Validation Panel

True to the new metaevaluation standard, a decision was made to again commission a Validation Panel whose purpose would be to perform that metaevaluation function during the development of the revised *Standards*. The revised *Standards* were to expand the original focus K-12 setting to now include higher education, professional education, business and industry, government, and the military. Panel candidates were to be nominated by the member organizations[1] and six individuals selected by the Joint

[1] The fifteen member organizations at the time *The Program Evaluation Standards* were developed were as follows: American Association of School Administrators, American Educational Research Association, American Evaluation Association, American Federation of Teachers, American Psychological Association, Association for Assessment in Counseling, Association for Supervision and Curriculum Development, Canadian Society for the Study of Education, Council of Chief State School Officers, Council on Postsecondary Accreditation, National Association of Elementary School Principals, National Association

Committee chair to reflect the expanded settings to which the *Standards* were to apply. A guest at the meeting, who had been a member of the Validation Panel for *The Personnel Evaluation Standards,* was nominated and appointed as chair of the new Validation Panel.

The Validation Panel was to: 1) identify and consider assumptions underlying the project; 2) critique and report on the Joint Committee's validation process using both the operating procedures approved by the American National Standards Institute and the new meta-evaluation standard; 3) assess the applicability of the revised program evaluation standards in various national and international contexts; 4) confront the committee with issues and ideas drawn from pertinent theoretical analyses and empirical research reports; and 5) report publicly on their assessment of the revised program evaluation standards.

The Validation Panel met with the Joint Committee in 1991, observed the development process, began its formative evaluation role and the following mission statement was proposed and adopted.

> The Validation Panel is charged with providing both formative and summative evaluation of the Joint Committee's process for revising *The Program Evaluation Standards* and the quality and utility of the final product. Specifically, the Panel will identify and consider assumptions underlying the project and critique and report on the Joint Committee's standards development process using the *Operating Procedures* approved by the American National Standards Institute (ANSI), the new metaevaluation standard, and applicable accepted professional practices. It will also assess the *Standards* applicability in national and international contexts, offer issues and ideas drawn from theoretical and empirical publications, and report publicly its assessment of the revised *Standards*. These duties will be discharged by 1) reviewing

---

of Secondary School Principals, National Council on Measurement in Education, National Education Association, National School Boards Association.

> the Joint Committee's plans, minutes, and position papers; and 2) observing the Committee's annual meetings and working groups, public hearings and field tests. Joint Committee, national reviewers, and field test products will be reviewed. Interim reports providing formative evaluation comments will be provided to the committee.

Validation Panel members were also present at subsequent Committee meetings in 1992 and 1993.

At each annual Committee meeting, beginning in 1991, the Panel provided written formative evaluation comments on the developmental process and the evolving standards. Additional areas of formative comments concerned methods for dealing with contributions of consultants, the panel of writers, the national review panel, the international review panel, field test participants, national hearings participants, and reviewers of the draft standards. The Panel comments are in the archived minutes of the Committee meetings.

The Standards

A brief description of the *Standards* is provided to provide the context for later comments. Thirty specific standards are categorized into groups relating to four attributes necessary and sufficient for sound and fair program evaluation -- utility, feasibility, propriety, and accuracy. Seven "utility standards are intended to ensure that an evaluation will serve the information needs of intended users." Three "feasibility standards are intended to ensure that an evaluation will be realistic, diplomatic, and frugal." Eight "propriety standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the

evaluation, as well as those affected by its results." And 12 "accuracy standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the program being evaluated." Each standard is accompanied by an overview of purpose, guidelines for applying the standard, examples of common errors and one or more illustrative cases and analysis descriptions. The format is virtually identical to that of *The Personnel Evaluation Standards* except that the *Personnel Standards* divided the overview into further sections (Explanation and Rationale) and contained a separate Suggestions section with each illustrative case. The 1994 *Standards* contain suggestions for applying the individual standards, including a checklist, and a tearout quick reference summary of the individual standards.

The Panel's Evaluation

The remainder of this report presents the Validation Panel's assessment of the standards project by considering the assumptions, the Joint Committee's validation process, and the applicability of the standards. The formative evaluation component of the Validation Panel's charge was earlier completed with submission of the last interim report after the Committee's 1993 annual meeting. Although the three interim reports were somewhat voluminous, a review of those annual reports revealed that in no case did the Committee fail to consider the issues raised even though addressing some issues required extensive effort. For example, at the Panel's suggestion the Committee categorized all reviewer, field test and hearings participants comments and documented

how they handled the issues raised. Publication of this final Panel report and presentations at national meetings will complete the Panel's requirement to report publicly on the revised *Program Evaluation Standards.*

ASSUMPTIONS

There are no explicit statements of guiding assumptions made in the *Standards* but there are numerous implicit assumptions as follows. First, there is a need for educational program evaluation standards. Secondly, the state-of-the-art in evaluation is such that experts can agree on the substance and format of such a set of standards. Thirdly, a single set of standards can be developed which meet the needs of all major stakeholders in the prescribed evaluation settings. Also, the 30 individual standards make implicit assumptions about the utility, feasibility, propriety, and accuracy of evaluations. Most of these later implicit assumptions are largely obvious. For example, the metaevaluation standard (A-12) assumes that evaluations guide program decisions, that flawed evaluations can mislead decision makers and that well done "evaluations of the evaluations" reduce the risk of significant flaws and add to the acceptability of the results.

The first assumption that there is a need for educational program evaluation standards is clearly valid. The 20,000 sales of the 1981 *Standards;* formal adoption of the *Standards* by at least one state legislature and informal reference by several legislatures; approval of the *Standards* by the American National Standards Institute (ANSI) as American National Standards; prevalence of the *Standards* as a primary text in university-level evaluation courses; enthusiastic support by hundreds of professional

evaluation educators and users of evaluations participants in the development, field test, and review process of the *Standards*; and the fact that the Validation Panel was unable to locate any competitive standards leaves little room for argument about this assumption's validity.

Concerning the second assumption, that the experts can agree on substance and format of a set of standards, they can and did agree as will be discussed in more detail concerning the Panel's second responsibility -- to critique the Joint Committee's validation process.

The third assumption that a comprehensive set of standards can be developed for use in all the target educational and training environments will also be dealt with in detail later. However, with only a small qualification, we can state that a comprehensive set of standards was developed for use in the targeted educational and training environments. That qualification is because the *Standards* readily state (page 2) that all standards do not apply equally to all environments and situations and that the user must exercise professional judgment to identify those that are applicable. The fact is that some users, e.g., parents, legislators, school board members and even novice evaluators may not have the background to make those judgments or realize when others have incorrectly made those judgments. The *Standards* have attempted to minimize this problem by providing a Functional Table of Contents which lists the "most relevant standards" for 10 frequent issues like deciding whether to evaluate, designing the evaluation, analyzing the information and reporting the results.

The individual implicit and explicit assumptions in the 30 specific standards and pro and con arguments will not be dealt with individually in this report. Those specific issues and arguments are well documented in the field test, review, and hearings comments in the Committee archives. All comments were reviewed by at least some Panel members and most Panel members were present when staff presented the aggregate comments and the Committee prescribed actions to take or endorsed staff recommendations. Clearly there were differences of opinion among external participants where polar views were expressed and the Committee had to take a position. For example, one consistent issue was whether the standards contained too much information and therefore would not be read and used by highly experienced evaluators or by macro-level decision makers like school boards. Others wanted extensive details so explicit guidance would be available when needed. The Committee opted to provide significant detail for those who might need the information but provided functional tables of contents, checklists and consolidated standards summaries for those who needed or wanted less detail.

VALIDITY OF THE STANDARDS DEVELOPMENT PROCESS

Introduction

This panel has evaluated the validity of the new *Standards* according to two criteria: First, the expertise and representativeness of the Joint Committee; and second, the process used to develop the standards. The issue is will the participants and process produce accepted practices and principles that will, if met, enhance the quality and

fairness of the practice of educational and training evaluation? In other words, is the process one which elicits the combined expertise of the appropriate participants, fairly considers divergent viewpoints and arrives at some majority or consensus viewpoint? Further, are those viewpoints packaged in such a fashion that users from varied backgrounds can apply them to evaluating the same training or educational program or materials and consistently arrive at similar conclusions? Therefore, this section will discuss the participants, the developmental process, and the validity of the product in that order.

Participants

Organizations represented. The 15 organizations who comprise the Joint Committee are the professional organizations which represent teachers, principals, administrators, counselors, school boards, accreditors, evaluators, curriculum developers, and researchers at all levels of education and training, public and private. Further, the Joint Committee administrative and oversight functions are provided by the Evaluation Center, Western Michigan University -- a noted and respected center of evaluation excellence.

The training and education stakeholders not directly represented by the organizations are the lay public, parents and legislators who pay for much of education and government training. Also absent are businesses and industry who receive the trained or educated students but many of their educators and trainers are members of the represented organizations. Although it is arguably the responsibility of superintendents

and school boards to look out for their interests, the fact is that parents and legislators were not directly represented. However, the Validation Panel membership did represent government, business and industry as well as many of the constituents of other Joint Committee member organizations. The National Parent Teacher Association did send a representative to the 1990 meeting but was unable to sustain support for subsequent working meetings.

The Validation Panel believes that the appropriate organizations were represented. Further, in 1989 the American National Standards Institute (ANSI) accredited the Joint Committee *Standards* developers as representing the national body of educational evaluation expertise. Those familiar with ANSI procedures will recognize this as an exacting and difficult procedure and reflects very favorably on the expertise represented in the Joint Committee.

Individual participants. With the exception of administrative staff, actual *Standards* development participants were all unpaid volunteers, many of whom donated hundreds of hours to the project. They volunteered by responding to their organizations newsletters, advertisements in journals and convention programs, word of mouth, or direct requests from the leadership of their respective organizations and associations. The credentials and representativeness of the participants are as impressive as their dedication to the project.

The participants represent a cast of many hundreds and a combined expertise of thousands of years in education and education and training evaluation. They served as Committee members, consultants, national review panel members, international review

panel members, conducted field tests, reviewers of multiple drafts, and participated in national hearings providing testimony and questioning those testifying. It is difficult to conceive of a viewpoint that was not represented to some degree.

In short, the Validation Panel believes that the level and representativeness of educational evaluation expertise could not be appreciably improved upon. Reaching out and involving people beyond those on the Joint Committee in drafting, testing and reviewing standards is consistent with the Committee's historical good-faith efforts to seek diverse points of view. Their efforts undoubtedly added breadth to the perspectives of the *Standards*. Although the panel of writers was drawn largely from the Committee member organizations, any lack of breadth there should have been compensated for by the diversity of the other players and the revision process which considered inputs from the many perspectives.

Development Process

Joint Committee Working Meetings. At each of the annual Committee working meetings, the Validation Panel was consistently impressed by the dedication and task orientation of the Committee and the manner in which input from all members of the Committee was fostered in a free exchange of ideas. Although *The Joint Committee By-Laws* (Article II, Section 5, MEETINGS) say that decisions are to be made by a simple majority vote, most decisions concerning the *Standards* content were the result of Committee consensus or near consensus. (New standards or significant change to an existing standard require a two-thirds vote.)

The basic procedure of the committee was to work in small focus groups to consider comments and changes suggested by external persons and draft changes. Validation Panel members rotated among the focus groups and made suggestions and comments but did not assume writing and revising responsibilities. The focus group decisions and revised material were then presented to the full Committee for additional change suggestions or approval. Again Panel members were present and free to offer comments and suggestions at the full Committee meetings. Like the Committee members, Panel members were made to feel that they were a part of the free exchange of ideas environment and their perspectives were seriously considered.

Interim reports of the Validation Panel. As previously stated, the Validation Panels comments and recommendations were summarized at the close of each annual meeting in 1991 through 1993 and were placed in the minutes. Not all members attended each meeting because of schedule conflicts and some limitation of resources for the 1992 and 1993 meetings but a quorum was always present.

The Joint Committee was responsive to the formative evaluations provided in the three interim reports. In no case were the criticisms or suggestions contained in the Validation Panel reports ignored.

Panel of Writers. As previously stated, the 1990 Committee meeting revised the 1981 original *Standards* by consolidating, revising and adding standards and making minor changes in the presentation format. Both the 1981 and 1994 Standards have 30 specific standards grouped under the headings of utility, feasibility, propriety, and accuracy although the wording and contents of some standards changed significantly.

Format changes were somewhat cosmetic, e.g., changing "pitfalls" to read "common errors" and deleting a "caveats" section. One major change was in the context of the illustrative cases to reflect the new orientation of the *Standards* as applicable to education and training efforts in settings that include schools, universities, law, medicine, nursing, the military, business, government, and social service agencies rather than the previous K-12 orientation.

In keeping with the Committee's commitment to obtain a broad range of input throughout the development process, a panel of writers was created to draft alternative versions of the new illustrative cases. Although the writers were largely recruited by the Committee member organizations, the diversity of the panel was a good fit to the expanded application settings of the new *Standards* and included 28 representatives from all major stakeholders except for students and legislators.

National and International Review Panels. In 1992, the first draft of the revised *Standards* was reviewed by 74 national and 14 international reviewers. These 88 reviewers provided a large volume of comments and again demonstrated the commitment of the Committee to have the standards reflect the collective judgments of as many view points as possible. Collating and summarizing the comments was a significant administrative challenge but the results were sufficient for the Committee to seriously consider and act on the reviews, resulting in many substantive changes. As described earlier, there were conflicting views to be considered and decisions made on which to incorporate in revisions. The Validation Panel commends the administrative staff on

their ability to consolidate and present the array of comments and the Committee for the diligence in considering each major point.

Attending to the advice of such a breadth of reviewers probably increased the acceptability and usefulness of the *Standards* and supports the notion of content validity. However, as in any such system of considering a range of views, those which are diverse and represent a minority view generally require more compelling evidence to be acted on favorably than more central views. Despite the intentions of the Committee, there is a chance given the magnitude of the data that minority comments could have been lost in the process. We have no evidence of any such loss, however.

Field Tests. Credible developers of tests, procedures, standards and training materials know that testing the products and procedures often averts disaster and identifies changes critical to their effective implementation. The Joint Committee is well aware of this principle and 32 professionals were recruited and in 1993 field tested the semi-final *Standards* and reported their judgments and recommendations. Despite a lack of fundin to commission field tests, the representative breadth of the settings is laudatory. In addtion to traditional school and university settings, legislative, municipal, business, social service, industry, and government agency settings were included.

Again the results were seriously considered by the Joint Committee, were used in making the final revisions to the *Standards*, and certainly contributed to the likelihood that they can be validly and efficiently applied in the desired settings. As in the original *Standards*, the new version appeals to users to share experiences in applying the *Standards* with the Joint Committee so that those experiences can be applied to the next

update of the *Standards*. Addresses for the experiences input and to obtain reporting forms are supplied in the book. Past experience in using the *Standards* was probably a significant contributor to the high proportion of successful field applications of the near-final *Standards*.

Public Hearings. Three public hearings were held during the Spring and early summer of 1993. The hearings were announced in newsletters and journals of the associations represented on the Joint Committee. They were timed to coincide with locations of national conventions frequented by educators and evaluators. Copies of the near-final *Standards* were sent to interested respondents and hearing times scheduled. Forty-six persons provided formal testimony at the hearings. Their orientations were significantly more divergent than would perhaps be expected by limiting advertising to the standard Joint Committee organizations and their journals but were not as diverse as say the field test settings. Nevertheless, the available advertisement opportunities did preclude some orientations, e.g., students and legislators, from learning about the hearings.

Again, the testimony was seriously considered by the Committee in making final changes to the *Standards*. Perhaps as a testimony to the comprehensiveness of the previous reviews and field tests, predominant testimony was very supportive of previous decisions, i.e., the *Standards* in their near-final format.

Comments on the Validity of the *Standards*

Developers of the *Standards* intended to provide guiding principles, not rules, for the conduct and use of educational and training program evaluations. How well did they succeed in fulfilling their intentions? *As* stated earlier, determining the validity of the *Standards* is a process of assessing the expertise and representativenes of the participants and determining if the developmental process permits a fair and open consideration of dominant and divergent expert viewpoints.

It is the judgment of the Validation Panel that the committee organized to revise the *Standards* was rich both in evaluation expertise and in representation from significant stakeholder groups. In addition, the process by which the committee carried out its work fostered an open exchange of ideas and views. The standards development process was essentially outlined by the Joint Committee starting in the Fall of 1975. The process has been evolving to become a mature and effective technique through continued use developing the 1981 *Program Evaluation Standards,* the 1988 *Personnel Evaluation Standards,* and now the 1994 *Program Evaluation Standards.*

The Validation Panel believes that the openness of the working meetings to hear and consider the ideas and views of their Committee members and others participating as writers, reviewers, field testers, testifiers at hearings and the Validation Panel members has produced a valid set of educational and training *Program Evaluation Standards* in the sense that they reflect current views of best practice and are responsive to the contextual concerns of various stakeholder groups. Similarly, the American National Standards Institute has accredited the Joint Committee and the 1994 *Standards* as the American

National Standard for educational program evaluation. The ANSI requirements process for accrediting representative groups and their standards is remarkably similar to the operating procedures used by the Joint Committee even before ANSI accreditation was a consideration.

Just as the standards developmental process has been evolving for the Joint Committee, so have three concepts central to educational evaluation: the qualitative/quantitative balance of evaluation, the orientation of validity, and the concept of metaevaluation.. The question here is, did the developmental process pick up on the evolution of the concepts? If so, this would be another indication of the validity of the process. Evaluation has progressed from a focus on quantitative summative program evaluation, with test score data the major indication of program success, to a combined formative and summative focus prevalent in the 1981 *Standards*. In the 1981 *Standards*, the emphasis was still on quantitative evaluation, with analysis focused on content, criterion-related, and construct validity as three separate components. In the interval between the 1981 and 1994 *Standards*, the old tripartite division of analytical approaches to evaluating test data has broken down; each approach is now seen as providing useful and mutually complementary information. Moreover, the concept of validity has come to include a more unified view, leading some to question the primacy or even the relevance of test score data in program evaluation. A more ethnographic or qualitative evaluation process began to emerge as a component of evaluation and created a lively debate among proponents of the polar ends of a qualitative/quantitative continuum. Clearly the value of qualitative evaluation and changed orientation of validity are reflected in the 1994

Standards. Along with the new metaevaluation standard, these changes probably represent the largest change in the 1994 over the 1981 *Standards*. Rogers (1994) and LaCompte (1994) provide detailed discussions of how this move to a unified view of validity is reflected in specific characteristics of the 1994 *Standards*. The central point here is that the developmental process used by the Joint Committee substantially incorporated these new views into the *Standards*, showing its responsiveness to the major intellectual trends in the field of measurement and evaluation.

APPLICABILITY OF THE STANDARDS

Purposes

The *Standards* are intended to

> provide a guide for evaluating educational and training programs, projects, and materials in a variety of settings. They are intended both for users of evaluations and for evaluators. People who commission or conduct evaluations, or who use evaluation results to improve education and training in schools, universities, medical and health care fields, the military, business and industry, the government, and law, will find the Standards useful. They have been developed for use by teachers, administrators, school board members, trainers, evaluators, curriculum specialists, legislators, personnel administrators, counselors, community leaders, business and educational associations, parents, and others. The Standards guide the design, employment, and assessment of evaluations of educational programs, projects, and materials (page 1).
>
> The Standards are not detailed technical standards, and they do not replace textbooks in technical areas such as qualitative and quantitative research design and analysis, measurement and data collection, data processing, and report writing (page 2).

Expert evaluators and stakeholders from each of the stated settings participated in development of the *Standards* and were specially targeted for inclusion in the review, field tests, and hearings processes. The illustrative cases specifically include each setting with the exception of the military, and that omission is not particularly important since the lessons of the cases generally apply equally well to the military formal training environment. Based on these points and the generally strong endorsement of the *Standards* by the external participants of the development process, the Validation Panel believes that the *Standards* claims of applicability and purpose are justified. However, we say this while recognizing that individual standards are not equally applicable in each setting and use. Further, as the *Standards* point out, other groups and settings who have not been included in the target uses may find additional uses for the standards. Clearly, the Committee was successful in expanding the utility of the 1994 *Standards* far beyond the target K-12 settings of the 1981 *Standards*.

One issue that is not clear is how applicable the *Standards* are for settings outside the United States? We expect that applicability in other countries would be variable but note that through both the ANSI and educational professional literature distribution, the availability and content of the *Standards* will be known throughout most of the developed world.

CONCLUSION

The Validation Panel confined its efforts to a charter to consider assumptions underlying the project, report on the validity of the developmental process, assess the

applicability of the standards, provide formative evaluation during the developmental process, and publicly report on the revised *Standards*.

The development and refinement process of the *Standards* was very systematic and open, incorporating diverse perspectives far beyond the general representation of the Joint Committee member organizations. The procedures and broad base of perspectives should enhance the acceptability of the *Standards* and likely resulted in a set of standards that represent the dominant state-of-the-art in educational program evaluation. Through its own charter, its formal commitments to ANSI and the Invitation to Users in the book, the Joint Committee is committed to a contined review and update of the *Standards* and as such the *Standards* can be viewed as a living set of principles. Recognition of the value of the *Standards* by the evaluation profession is already documented in reviews by Fournier (1994), Hanson (1994), and Patton (1994).

REFERENCES

Fournier, D. (1994). Review of *The Program Evaluation Standards. Journal of Educational Measurement, 31(4),* 363-367.

Hanson, J. (1994). Review of *The Program Evaluation Standards. Evaluation Practice, 15(2),* 189-92.

LaCompte, M. D. (1994) Sensible match-making: Qualitative research design and *The Program Evaluation Standards. Journal of Experimental Education, 63(1),* 29-44.

Linn, R. E., Buchmann, M., Gould, R. B., Kellaghan, T., Lawrence, D., Robinson, P., & Zirkel, P. (1988). Development, validation, and application of the personnel evaluation standards. *Journal of Personnel Evaluation in Education, 2,* 199-214.

Patton, M. Q. (1994). Review of *The Program Evaluation Standards. Evaluation Practice, 15(2),* 193-199.

Rogers, W. T. (1994). The treatment of measurement issues in the revised *Program Evaluation Standards. Journal of Experimental Education, 63(1),* 13-28.

The Joint Committee on Standards for Educational Evaluation. (1994). *The Program evaluation standards: How to assess evaluations of educational programs.* Thousands Oaks, CA: Sage Publications.

The Joint Committee on Standards for Educational Evaluations. (1988). *The Personnel evaluation standards: How to assess systems for evaluating educators.* Thousands Oaks, CA: Sage Publications.

The Joint Committee on Standards for Educational Evaluations. (1981). *Standards for evaluations of educational programs, projects, and materials.* New York, NY: McGraw-Hill Book Company.